Formal methods for risk analysis of privacy-preserving data publishing algorithms

M2 thesis proposal

Supervisors: Tristan Allard (Univ. Rennes, IRISA), Barbara Fila (INSA Rennes, IRISA)

Contact: tristan.allard@irisa.fr and barbara.fila@irisa.fr

Keywords: privacy-preserving data publishing, formal methods, privacy attacks, risk analysis.

Context

Large volumes of personal data are nowadays collected by private companies or public organizations. Typical examples include health records, geolocation, electricity consumption, or social networks. Various legal[1], monetary[2], or visibility incentives encourage data holders to share anonymized versions of the collected datasets. On the one hand, data publishing is useful because it allows to strengthen scientific studies, to favor industrial innovation, or to establish appropriate public policies. Numerous open data repositories are well known today[3] and their number keeps growing[4]. On the other hand, publishing personal data must be done with care in order to provide strong privacy guarantees. Failing to safeguard privacy may lead to severe consequences on individuals[5] or companies[6]. To preserve privacy, privacy-preserving data publishing techniques are used. Today's typical techniques include *ad-hoc* aggregation[7], k-anonymous generalization [5, 6], or differentially private perturbation [1]. Their aim is to ensure that the published data enjoy (hopefully strong enough) privacy guarantees. Security is however a constant race between the attackers and the defenders. A large number of attacks on privacy-preserving data publishing algorithms exists today and keeps growing [4].

Objective

The objective of this project is to develop a formal framework (see, *e.g.,* [8]) allowing to analyze risks relevant to privacy-preserving data publishing (see, *e.g.,* [7]).

The tasks of the student will be to:

- Perform a systematic analysis of the state of the art on existing privacy-preserving data publishing techniques and related attacks.
- Analyze existing attacks in order to identify aspects and/or parameters (*e.g.,* privacy loss, usability, attacker's background knowledge) that should be taken into account while analyzing the risks relevant to a privacy-preserving data publishing technique.
- Propose a taxonomy of relevant attackers, based on aspects such as adversarial goals, background knowledge, reasoning methods, computational capabilities, *etc.*
- Develop a model allowing to evaluate privacy risks in the personal data publishing context.

[1] See for example the EU Directive 2019/1024 on open data and the re-use of public sector information.

[2] For example, through data marketplaces such as Innodata (https://innodata.com/ai-data-marketplace/) or Defined.ai (https://www.defined.ai/).

[3] See for example the French open data repository (https://www.data.gouv.fr/fr/), or private initiatives like the COVID-19 Google Health repository (https://health.google.com/covid-19/open-data/).

[4] See, for example, the open data portal of Rennes Métropole https://blog.rudi.bzh/

[5] See for example the 23andme recent breach (https://www.theguardian.com/technology/2024/feb/15/23andme-hack-data-genetic-data-selling-response).

[6] See for example the lawsuit following the 23andme breach (https://www.theverge.com/2024/9/13/24243986/23andme-settlement-dna-data-breach-lawsuit).

[7] See for example the data published by Enedis, the French electricity distribution system operator (https://data.enedis.fr/explore/?source=shared&sort=modified).

Supervision

This master project is proposed by the Security and Privacy (SPICY) team from the IRISA institute in Rennes, France. The work will be supervised jointly by Tristan Allard (PhD, HDR) associate professor at the University of Rennes, expert in privacy in data intensive systems, and Barbara FILA (PhD, HDR), associate professor at INSA Rennes, expert in formal methods for risk assessment.

Candidate profile and application

We are looking for a candidate who is:

- interested in security and privacy,
- enthusiastic about formal modeling,
- curious and open-minded,
- speaking English or French (knowledge of French is not mandatory).

If the results of this master project are satisfactory, pursuing on a PhD position will be considered.

To apply, please send the following documents to both `tristan.allard@irisa.fr` and `barbara.fila@irisa.fr`:

- your CV,
- the grade transcript of all university-level courses taken.

Please contact `tristan.allard@irisa.fr` and `barbara.fila@irisa.fr` if you have questions. Informal inquiries are welcome.

References

[1] Cynthia Dwork. Differential privacy. In *Proceedings of the 33rd International Conference on Automata, Languages and Programming - Volume Part II*, ICALP'06, pages 1–12, 2006.

[2] Barbara Fila and Wojciech Widel. Efficient attack-defense tree analysis using pareto attribute domains. In *CSF*, pages 200–215. IEEE, 2019.

[3] Hongsheng Hu, Zoran A. Salcic, Lichao Sun, Gillian Dobbie, P. Yu, and Xuyun Zhang. Membership inference attacks on machine learning: A survey. *ACM Computing Surveys (CSUR)*, 54:1 – 37, 2021.

[4] Ahmed Salem, Giovanni Cherubin, David Evans, Boris Köpf, Andrew Paverd, Anshuman Suri, Shruti Tople, and Santiago Zanella-Béguelin. Sok: Let the privacy games begin! a unified treatment of data inference privacy in machine learning. In *Proceedings of the 2023 IEEE Symposium on Security and Privacy (S&P '23)*, pages 327–345, 2023.

[5] Pierangela Samarati. Protecting respondents identities in microdata release. *IEEE Transactions on Knowledge and Data Engineering*, 13(6):1010–1027, 2001.

[6] Latanya Sweeney. k-anonymity: a model for protecting privacy. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, 10(5):557–570, 2002.

[7] Antonin Voyez, Tristan Allard, Gildas Avoine, Pierre Cauchois, Élisa Fromont, and Matthieu Simonin. Membership inference attacks on aggregated time series with linear programming. In *Proceedings of the 19th International Conference on Security and Cryptography (SECRYPT '22)*, 2022.

[8] Wojciech Widel, Maxime Audinot, Barbara Fila, and Sophie Pinchinat. Beyond 2014: Formal methods for attack tree-based security modeling. *ACM Comput. Surv.*, 52(4):75:1–75:36, 2019.